

ON SOME COMMON INTERESTS AMONG RELIABILITY,
INVENTORY AND QUEUING

by

Donald Gross

GWU/IMSE/Serial T-491/84
13 June 1984

The George Washington University
School of Engineering and Applied Science
Institute for Management Science and Engineering

Research Supported by
Contract N00014-83-K-0217
Office of Naval Research
and
Grant ECS-8200837-01
National Science Foundation

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

Abstract
of
GWU/IMSE/Serial T-491/84
13 June 1984

ON SOME COMMON INTERESTS AMONG RELIABILITY,
INVENTORY AND QUEUING

by

Donald Gross

Queuing networks can be used to model maintained systems. Under many conditions, closed queuing network theory can be applied to ascertain the availability of such systems. Multi-echelon repairable item inventory systems serve as one such class of examples. Problems of common interest to the reliability, queuing, and inventory communities are highlighted, and solution techniques for these problems presented.

Research Supported by
Office of Naval Research
and
National Science Foundation

THE GEORGE WASHINGTON UNIVERSITY
School of Engineering and Applied Science
Institute for Management Science and Engineering

ON SOME COMMON INTERESTS AMONG RELIABILITY,
INVENTORY AND QUEUEING

by

Donald Gross

1. Introduction

The major purpose of this paper is to illustrate a class of problems which are of mutual interest to the reliability, queuing and inventory communities. Although often separately studied, interests in such problems really are common, and mutual benefits could accrue by interaction among these communities.

2. A Reliability Problem

In Mann, Schafer and Singpurwalla (1974), Section 10.3 deals with reliability models for maintained systems. In particular, Section 10.3.1 gives an example of a single unit which fails according to an exponential distribution with mean time to failure (MTTF) of, say, $1/\lambda$ and is repaired (as good as new) according to an exponential distribution with mean time to repair (MTTR) of, say, $1/\mu$. This process is then a continuous time Markov process (CTMP) and is driven by the

infinitesimal generator or, as it is also often called, the rate matrix

$$Q = \begin{matrix} & 0 & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \end{matrix} . \quad (1)$$

The two possible system states are 0 (unit is operating) or 1 (unit is down, and undergoing repair).

We desire to find the availability of the unit at time t , which we denote as $A(t)$, and to do this we need to find $p(t)$, the state probability (row) vector at time t , that is,

$$p(t) = (p_0(t), p_1(t))$$

and hence

$$A(t) \equiv p_0(t) .$$

To find $p(t)$, we must solve the Kolmogorov forward equations (a set of differential-difference equations)

$$p'(t) = p(t)Q , \quad (2)$$

with the added condition that the probabilities sum to one, namely,

$$1 = p(t)e , \quad (3)$$

where e is a column vector of 1's. Thus writing out (2) and (3) we have

$$p'_0(t) = -\lambda p_0(t) + \mu p_1(t) \quad (4)$$

$$p'_1(t) = \lambda p_0(t) - \mu p_1(t) \quad (5)$$

$$1 = p_0(t) + p_1(t) , \quad (6)$$

and we must solve the set of equations (4) and (6) or (5) and (6). This can be easily done using Laplace transforms [we employ the boundary

condition $(p_0(0) = 1, p_1(0) = 0)$; that is, the unit is working at time zero] and obtain

$$A(t) = p_0(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} . \quad (7)$$

Note that the steady state availability is

$$A \equiv \lim_{t \rightarrow \infty} A(t) = \frac{\mu}{\lambda + \mu} ,$$

the well known result for an alternating renewal process.

3. An Expanded Reliability Problem

We now consider an expanded version of the problem treated in Section 2. Consider now N units and c repair channels ($c \leq N$) . We now define $A(t)$ as the probability that at least some desired number, say M , of the units is operational at time t . If more than M are operational, the excess are considered spares and are on cold standby (note that there are a total of $N - M \equiv y$ spares in the system, but that all y spares are not always available). If less than M units are operational, the system is performing below the desired level.

A system state can be described by the number of units up (or operating, call this n_U) or by the number of units in or awaiting repair (call this n_R) . Either state descriptor gives complete information since $n_U + n_R = N$. For this problem, the Q matrix is $N+1 \times N+1$ as there are a total of $N+1$ states: $0, 1, 2, \dots, N$. Hence it is necessary to solve a set of $N+1$ linear, first-order differential equations of the type given by (2).

4. A Queuing Problem

The above "reliability problem" is also a "classical" problem in queuing theory and is known as the machine repair problem [see Cooper (1981, Section 3.8), Kleinrock (1974, Section 3.8), or Gross and Harris (1974, Section 3.6)]. Figure 1 shows a schematic of this problem, modeled as a closed queuing network. This is a two node, closed queuing network, where the total of N units are, at various times and in various combinations, distributed among the two nodes. At the operating node, we show M parallel service channels so that a queue at this node represents the cold standby spares available.

5. A Repairable Item Inventory Problem

The problem discussed in Sections 3 and 4 also fits the category of an inventory problem. It is a typical "repairable item inventory problem," for which it is desired to find the optimal combination of the numbers of spares and repair channels, so as to satisfy certain service

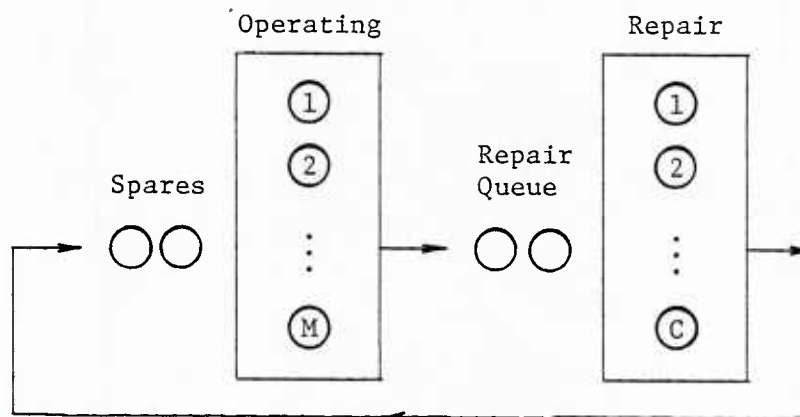


Figure 1. Schematic of a machine repair problem.

level performance criteria. Thus the problem mathematically is to find y and c which

$$\text{Minimize } E[\text{Cost/Year}] = k_1 y + k_2 c$$

$$\text{subject to } A(t_i) \equiv \sum_{n=0}^y p_n(t_i) \geq 1 - \alpha \quad (i = 1, 2, \dots, T)$$

$$L(t_i) \equiv \sum_{n=0}^N n p_n(t_i) \leq \mathcal{L} \quad (i = 1, 2, \dots, T) ,$$

⋮

where p_n is the probability that n units are at the repair node, k_1 is the annual cost associated with having a spare (amortization of purchase cost including interest, insurance, storage, etc.), k_2 is the annual cost associated with each repair channel (amortization, salary of repair crew, maintenance of repair equipment, etc.), $1 - \alpha$ is the desired availability, and \mathcal{L} is the desired limit on the average number of units in or awaiting repair. The dots represent other constraints that may possibly be imposed, for example, a constraint on total budget.

In order to solve this problem, it is necessary first to find $p(t)$, and this is what we focus our attention on here. We refer the reader to Gross, Miller and Soland (1983) for a discussion of the optimization aspects of such problems.

Consider a more complex multi-echelon version of the above problem as shown in Figure 2. Pictured here are three "field" locations, each with local repair capability. However, depending on the problem causing the failure, a certain percentage of failed units must be sent

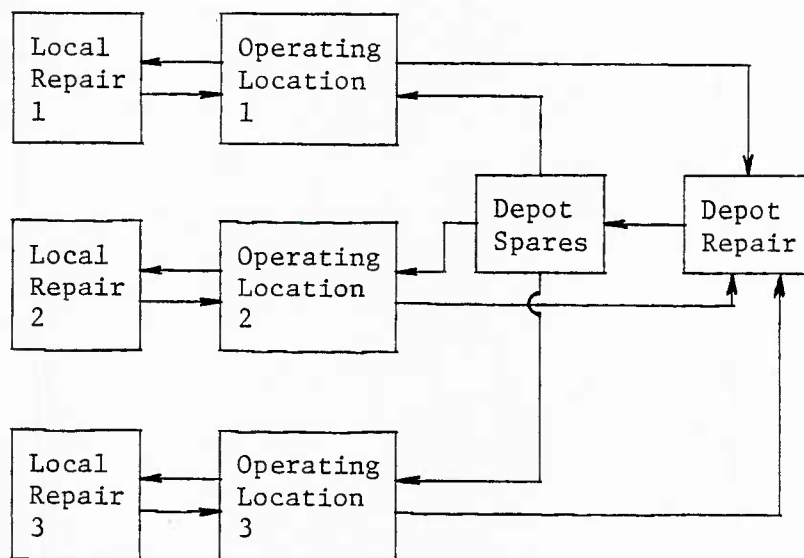


Figure 2. Multi-echelon repairable item inventory system.

to a higher echelon (depot) to be repaired. Each field location, as well as the depot, stocks spare units which, if available, are dispatched from the location to which the failed unit is sent. If spares are not available, requests are backordered.

As long as all failure and repair times are exponential, we still have a CTMP, albeit with a very large (but finite) Q matrix. For example, Table 1 shows a specific example which yields a state space of over 100 million states.

For such systems, shown in Figure 2, we might desire $A_1(t)$, $A_2(t)$, $A_3(t)$ and $A_{123}(t)$, where $A_i(t)$ is the probability that M_i units are operating at field location i at time t ($i = 1, 2, 3$) and $A_{123}(t)$ is the probability that at time t , M_1 are operating at location 1, M_2 at location 2, and M_3 at location 3 simultaneously. In the example given in Table 1, $M_i = 25$, $i = 1, 2, 3$.

TABLE 1
THREE LOCATION, TWO ECHELON EXAMPLE

Location	N	M	c
1	29	25	3
2	29	25	3
3	29	25	3
Depot	6	--	4

Number of states = $|\mathcal{S}| = 100,706,625$

6. Solution Techniques

Obviously, for large systems, the use of Laplace transforms for obtaining $p(t)$ is not feasible. Since our systems are finite, numerical methods can be utilized. Numerical integration techniques such as Runge-Kutta or predictor-corrector can be employed for moderately sized systems. We found for these types of problems, another method which we refer to as *randomization* to be more efficient. For details of the development of this procedure, we refer the reader to Grassmann (1977a and 1977b), or to Gross and Miller (1984a, 1984b). The randomization method, as far as we can ascertain, dates back at least to a paper by Jensen (1953), and is mentioned, often under other names (for example, subordination of Markov chains to Poisson processes or uniformized embedded Markov chains), by Cohen (1969), Feller (1971), Çinlar (1975), Keilson and Kester (1977), and Keilson (1979), to mention a few.

The basic idea of the randomization technique is to view the CTMP in a certain way, which allows the major computation to be performed on an imbedded discrete time Markov chain (DTMC) called the *uniformized* chain. The transitions for this DTMC are generated by an underlying Poisson process (hence the name randomization). The single-step transition probability matrix of the DTMC and the parameter (rate) of the Poisson process are functions of the original rate matrix, $Q = \{q_{ij}\}$, of the CTMP.

Let

$$\Lambda = \max_i |q_{ii}|$$

and

$$P = Q/\Lambda + I.$$

Then the imbedded uniformized DTMC has single-step transition probability matrix P and the transitions of this DTMC are generated by a Poisson process with rate Λ . Note that since the diagonal elements of Q are negative, that is,

$$q_{ii} = - \sum_{\substack{j \\ (i \neq j)}} q_{ij},$$

Λ is actually the absolute value of the minimum diagonal element, which is the mean exit rate of the state with the largest mean exit rate.

Denoting the state probability vector after k transitions of the DTMC by $\phi(k)$, it can be shown [see Gross and Miller (1984a)] that

$$p_j(t) = \sum_{k=0}^{\infty} \phi_j(k) \frac{(\Lambda t)^k e^{-\Lambda t}}{k!}, \quad (8)$$

where $p_j(t)$ is the probability that the CTMP is in state j at time

t (jth element of $\underline{p}(t)$), $\phi_j(k)$ is the probability that the imbedded uniformized DTMC is in state j after k transitions (jth element of $\underline{\phi}(k)$) and $(\Lambda t)^k e^{-\Lambda t}/k!$ is the probability of k transitions of the DTMC in clock time t . The usual recursion can be used to get $\phi_j(k)$, that is,

$$\underline{\phi}(0) = \underline{p}(0) ; \quad \underline{\phi}(k) = \underline{\phi}(k-1)P . \quad (9)$$

To use (8) for computational purposes, the infinite sum must be truncated. The error of truncation can be nicely bounded since we are discarding a Poisson tail; and, in fact, the computing version of (8) becomes

$$p_j(t) = \sum_{k=0}^{T(\epsilon, t)} \phi_j(k) \frac{(\Lambda t)^k e^{-\Lambda t}}{k!} , \quad (10)$$

where

$$\sum_{k=0}^{T(\epsilon, t)} \frac{(\Lambda t)^k e^{-\Lambda t}}{k!} \geq 1 - \epsilon , \quad (11)$$

ϵ being the desired error bound. One advantage of this method over numerical integration (besides efficiency) is the ability to exactly bound the computational error.

7. Results

The largest problem solved directly by the procedure described in the previous section is shown in Table 2. This example is a two field location, two echelon system with a state space size of 20,748. Calculated were $A_1(t)$, $A_2(t)$, $A_{12}(t)$, $t = 1, 2, \dots, 15$, with the following time-varying scenario. At time $t = 6$, a sudden decrease

TABLE 2
TWO LOCATION, TWO
ECHELON EXAMPLE

Location	N	y	c
1	18	4	2
2	13	2	2
Depot	--	3	4

$$\epsilon = .001$$

$$|\$| = 20,748$$

of MTTF occurs. The repair facilities cannot make an "in kind" accommodation until time 10 . Figure 3 shows a plot of $A_1(t)$ versus t [$A_2(t)$ and $A_{12}(t)$ are similar in nature]. The graph shows an initial $A_1(0)$ of 1.0 (we assume at time zero *all* units are operational) and thereafter a drop-off toward the steady-state availability as time increases. At time 6 , the increase in failure rate occurs and $A(t)$ begins to drop off at an increasing rate, heading for a new, lower steady-state availability. However, the increase in repair rate at time 10 causes $A(t)$ to begin to rise, heading back toward the original steady-state availability.

This run took approximately 25 minutes of CPU time on a VAX 11/780 computer using the randomization computation of (10) with a more efficient procedure than the recursion of (9) [given in Gross and Miller (1984a)] for calculating $\phi(k)$.

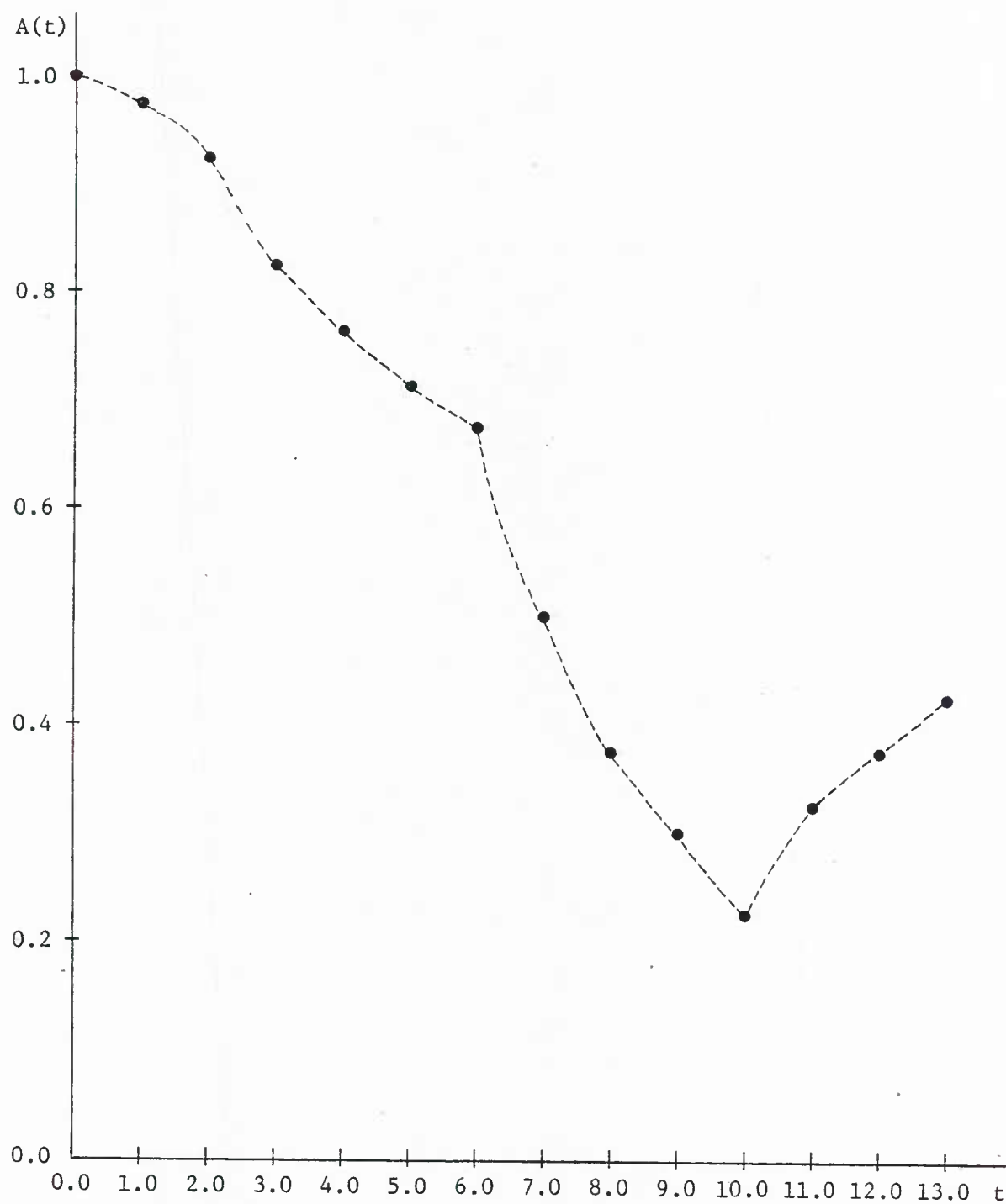


Figure 3. $A_1(t)$ versus t for sample run.

As the systems become more complex (more bases, multiple component types, indenture, more echelons, etc.) the state-space grows rapidly. We have solved a three location problem, shown in Table 3, using a truncated state-space approach, where seldom visited states are "lumped" together in single absorbing states [see Gross, Kioussis, and Miller (1984)]. There are over 43 million states, but via the truncation approach, the state-space was reduced to 23,410 and solved in about 30 minutes CPU time on the VAX 11/780, adding an error of .007. Calculated were $A_1(t)$, $A_2(t)$, $A_3(t)$ and $A_{123}(t)$, for $t = 1, 2, \dots, 15$.

TABLE 3
THREE LOCATION, TWO
ECHELON EXAMPLE

Location	N	y	c
1	25	1	1
2	24	1	1
3	24	1	1
Depot	--	2	1

$$\epsilon = .001$$

$$|S| = 43,278,703$$

8. Conclusions

We have presented here a class of problems of interest to the reliability, queuing and inventory communities and briefly demonstrated a viable solution procedure for these problems. While researchers in the above communities often go their "separate ways," better communication among them should benefit all.

9. Acknowledgments

Acknowledgment and thanks are due to my coworkers at George Washington University, Professors D. R. Miller and R. M. Soland, and to Mr. L. C. Kioussis, who have made enormous contributions to the solution of the problems discussed in this paper.

10. References

- GINLAR, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, New Jersey.
- COHEN, J. W. (1969). *The Single Server Queue*. North Holland, Amsterdam.
- COOPER, R. B. (1981). *Introduction to Queueing Theory*, 2 Ed. North Holland, Amsterdam.
- FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications, Vol. II*, 2 Ed. Wiley, New York.
- GRASSMANN, W. (1977a). Transient Solutions in Markovian Queueing Systems, *Computers & Operations Research*, 4, 47-56.
- GRASSMANN, W. (1977b). Transient Solutions in Markovian Queues, *European Journal of Operational Research*, 1, 396-492.
- GROSS, D. and C. M. HARRIS (1974). *Fundamentals of Queueing Theory*. Wiley, New York.
- GROSS, D., L. C. KIOUSSIS, and D. R. MILLER (1984). Transient Behavior of Large Markovian Multi-echelon Repairable Item Inventory Systems Using a Truncated State Space Approach, Technical Paper GWU/IMSE/T-488/84, Institute for Management Science and Engineering, George Washington University.
- GROSS, D. and D. R. MILLER (1984a). The Randomization Technique as a Modeling Tool and Solution Procedure for Transient Markov Processes, *Operations Research*, 32, 343-361.
- GROSS, D. and D. R. MILLER (1984b). Multi-echelon Repairable Item Provisioning in a Time-varying Environment Using the Randomization Technique, *Naval Research Logistics Quarterly*, in press.

- GROSS, D., D. R. MILLER, and R. M. SOLAND (1983). A Closed Queuing Network Model for Multi-echelon Repairable Item Provisioning, *IEEE Transactions*, 15, 344-352.
- KEILSON, J. (1979). *Markov Chain Models—Rarity and Exponentiality*. Springer-Verlag, New York.
- KEILSON, J. and A. KESTER (1977). Monotone Matrices and Monotone Markov Processes, *Stochastic Processes and Their Applications*, 5, 231-241.
- KLEINROCK, L. (1975). *Queueing Systems, Vol. I*. Wiley, New York.
- MANN, N. R., R. E. SCHAFER, and N. D. SINGPURWALLA (1974). *Methods of Statistical Analysis of Reliability and Life Data*. Wiley, New York.

U213993
U 213907

THE GEORGE WASHINGTON UNIVERSITY

BENEATH THIS PLAQUE
IS BURIED

A VAULT FOR THE FUTURE
IN THE YEAR 2056

THE STORY OF ENGINEERING IN THIS YEAR OF THE PLACING OF THE VAULT AND
ENGINEERING HOPES FOR THE TOMORROWS AS WRITTEN IN THE RECORDS OF THE
FOLLOWING GOVERNMENTAL AND PROFESSIONAL ENGINEERING ORGANIZATIONS AND
THOSE OF THIS GEORGE WASHINGTON UNIVERSITY

BOARD OF COMMISSIONERS, DISTRICT OF COLUMBIA
UNITED STATES ATOMIC ENERGY COMMISSION
DEPARTMENT OF THE ARMY, UNITED STATES OF AMERICA
DEPARTMENT OF THE NAVY, UNITED STATES OF AMERICA
DEPARTMENT OF THE AIR FORCE, UNITED STATES OF AMERICA
NATIONAL ADVISORY COMMITTEE FOR AERONAUTICS
NATIONAL BUREAU OF STANDARDS, U.S. DEPARTMENT OF COMMERCE
AMERICAN SOCIETY OF CIVIL ENGINEERS
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS
THE SOCIETY OF AMERICAN MILITARY ENGINEERS
AMERICAN INSTITUTE OF MINING & METALLURGICAL ENGINEERS
DISTRICT OF COLUMBIA SOCIETY OF PROFESSIONAL ENGINEERS, INC.
THE INSTITUTE OF RADIO ENGINEERS, INC.
THE CHEMICAL ENGINEERS CLUB OF WASHINGTON
WASHINGTON SOCIETY OF ENGINEERS
FAULKNER, KINGSBURY & STENHOUSE - ARCHITECTS
CHARLES H. TOMPKINS COMPANY - BUILDERS
SOCIETY OF WOMEN ENGINEERS
NATIONAL ACADEMY OF SCIENCES, NATIONAL RESEARCH COUNCIL

THE PURPOSE OF THIS VAULT IS INSPIRED BY AND IS DEDICATED TO
CHARLES HOOK TOMPKINS, DOCTOR OF ENGINEERING
BECAUSE OF HIS ENGINEERING CONTRIBUTIONS TO THIS UNIVERSITY, TO HIS
COMMUNITY, TO HIS NATION, AND TO OTHER NATIONS.

BY THE GEORGE WASHINGTON UNIVERSITY.

ROBERT V. FLEMING
CHAIRMAN OF THE BOARD OF TRUSTEES

CLOYD H. MARVIN
PRESIDENT

JUNE THE TWENTIETH
1956

To cope with the expanding technology, our society must be assured of a continuing supply of rigorously trained and educated engineers. The School of Engineering and Applied Science is completely committed to this objective.